

1.1.2 – Defining the operating environment

Practical guidance – healthcare

Author: SAM demonstrator project

Healthcare is a complex and diverse setting with many different operating environments. A family doctor's practice is very different from a hospital setting, and even within a hospital there is diversity across operating environments such as surgery or the hospital pharmacy. Reflecting this broad range of potential operating environments is the large number of different types of artificial intelligence (AI) and machine learning (ML) applications in healthcare. Examples include clinician-facing applications (e.g. breast cancer screening algorithms), patient-facing mobile phone apps (e.g. symptom checkers), and tools to support healthcare business processes (e.g. missed appointment predictors).

The definition of the operating environment can, therefore, be challenging for developers of AI and ML applications in healthcare. Drawing an accurate boundary around the AI/ML system and the operating environment is not straightforward, and can be done in different ways. To date, most developers have bounded the AI/ML system very narrowly and assumed a well-defined task or function in order to reduce complexity. For example, one way of looking at an algorithm for breast cancer screening is to consider only a set of mammograms as input and the likelihood of malignancy as the output. However, this approach runs into difficulties quickly when the wider use context needs to be considered, for example when an algorithm trained on data from a specific population or health system (e.g. patients in the NHS in the UK) is deployed in another population or health system (e.g. patients in the US). Performance figures tend to drop quickly in these situations.

Another option is to define the operating environment as the clinical system within which the AI/ML will be used. This perspective recognises that the AI/ML interacts with other technology and with people. Care is generally delivered by teams of healthcare professionals working as clinical teams, and supported by a large number of tools and technologies. AI and ML systems, even with increasing autonomy, might be best understood as part of such clinical teams.

A useful approach to model clinical systems at the functional level is the Functional Resonance Analysis Method (FRAM) [1]. FRAM decomposes the clinical system into functions, to move away from “what a system is” to “what it does”. Each function is examined for its potential performance variability, then interactions between functions are examined. “Functional resonance” is used to describe how outcomes can “emerge” from everyday variability of many functions, to move away from simple notions of “cause and effect”.

FRAM is built on four principles:

1. **The principle of equivalence of success and failure** – success and failure come from the same source, i.e. they are not fundamentally different in nature. Approximate adjustments mean that people adapt successfully most of the time but sometimes variability in performance will lead to unsatisfactory outcomes.

2. **The principle of approximate adjustments** – due to limitations in resource, uncertainties, underspecified systems and variance demands people will adjust to suit the situation. This gives rise to performance variability which is inevitable, ubiquitous and necessary.
3. **The principle of emergence** – complex systems with many links and fluctuating approximate adjustments become intractable as it is impossible to predict what will happen precisely beyond expecting regular events.
4. **The principle of functional resonance** – functions represent the different things a system does. Due to approximate adjustments these will exhibit performance variability. Functional resonance refers to how functions may impact each other's performance variability. Small changes could lead to disproportionately large effects and vice versa.

The strength of FRAM is that it supports the analyst or system designer in reasoning about interactions. For example, when introducing an autonomous infusion pump into the intensive care unit, FRAM encourages consideration of not just the algorithmic performance (e.g. whether the infusion pump can control a patient's blood sugar levels by giving insulin), but also of how the autonomous infusion pump communicates with nurses and doctors as well as other systems, such as the electronic patient record. This provides a more realistic representation of the complexity of the operational environment in healthcare settings.

References

[1] Hollnagel E. FRAM, the functional resonance analysis method: modelling complex socio-technical systems: Ashgate Publishing, Ltd.; 2012.